Toward a Comprehensive Bottom-Up and Top-Down Analysis of a Reference Monoclonal Antibody

Chris Becker[1] Wilfred Tang[1] Yong J. Kil[1] Marshall Bern[1] John Schiel [2] Lisa Kilpatrick[2] Trina Formolo[2]
[1]Protein Metrics Inc., San Carlos, CA, USA [2]National Institutes of Standards and Technology (NIST) Gaithersburg, MD, USA Contact: becker@proteinmetrics.com

Abstract

Therapeutic proteins now command a major share of new medicines, and monoclonal antibodies (mAbs) are the most common in this class because they can be readily engineered for a wide range of disease targets. However, careful characterization is a difficult task and is required due to their inherent variability. Mass spectrometry is a leading tool for this characterization because of its ability to analyze all kinds of modifications as well as verify the primary sequence. In this study, high resolution mass spectrometric data of the NIST mAb reference material was analyzed by the new bioinformatics tools Byonic and Byologic to identify and quantify a wide variety of modifications, including oxidation, glycation, glycosylation, and sequence variants. Examples of modifications include but are not limited to:

- Sequence Variant Analysis (SVA)
- Oxidation
- Deamidation, ammonia loss
- Disulfide-bonded peptides
- Glycation and Glycosylation
- Antibody-drug conjugates
- Truncations and Extensions
- Unanticipated Modifications

Prior and New Workflow

A number of different other tools exist to perform analyses such as sequence variant analysis. However they are fragmented, inefficient, and raise the risk of missing sequence variants or reporting false positives.

Prior Workflow:

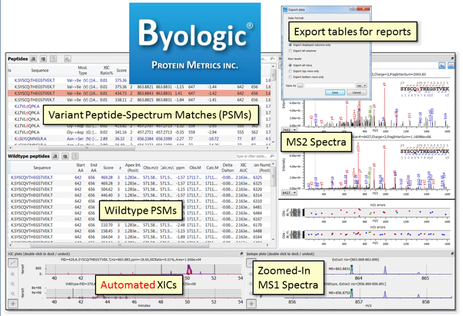
1. Many separate tools
2. Database search engine for MS2
3. Sequence editor/viewer
4. Vendor MS software for manual XIC
5. Scripts and/or Excel for area calculations and report generation
6. Long project times

New Workflow: Byonic + Byologic

Current Method

Heavy and light chain components were measured by both bottom-up and top-down (sometimes called middle-down) approaches by a high resolution Thermo Fisher Orbitrap Elite mass spectrometer employing CID, HCD and ETD fragmentation modes. Bottom-up digestion was by trypsin, chymotrypsin and glu-C. Intact ~ 25 kDa parts of the reference mAb, light chain and IdeS-digested heavy chain components Fd' and scFc, were analyzed in the top-down assay. The data was analyzed by a combination of the search engine Byonic™ (Bern et al. 2012) and new inspection software Byologic® that combines and compares MS1 and MS2 data streams and performs label-free quantification by taking the ratio of extracted ion chromatograms (XICs) of the modified to unmodified peptide. In addition, new peptide mapping software, Byomap™, was employed to quantify and annotate the peptide map (not shown due to space limitation).

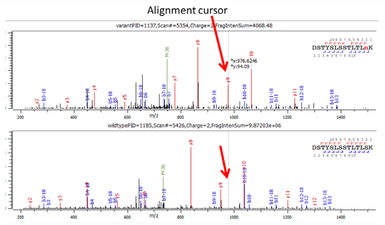
Overview. Examples: sequence coverage, and sequence variants



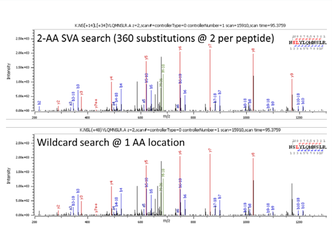
Overview of the dashboard of Byologic: inspection and quantification software

Primary sequence coverage, using three different digestive enzymes.

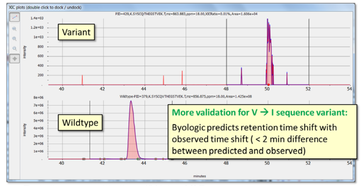
Sequence partially redacted sequence until release of the reference material.



Side-by-side comparison of the modified and unmodified annotated MS2 spectra of a sequence variant.



Peptide with two AA substitutions also detected in test: NSLYLQMNSLR → NTFYLQMNSLR, in 1% mAb spike by two independent, complementary approaches

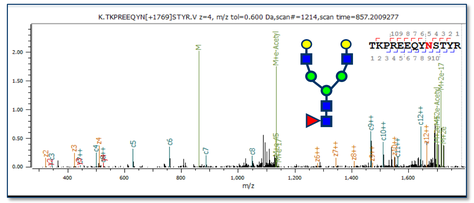


Relative quantitation (%) by XIC areas: Variant / (Variant + Wildtype) x 100 = 0.01%. Can check and if necessary adjust XIC integration window.

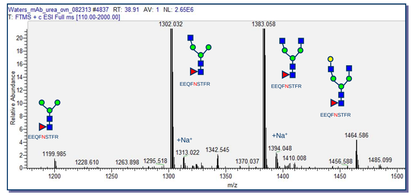
XIC Ratio%	XIC AUC	Sequence	Mod. Type	MS res.
0.049	1.32E+05	K.STSGGTAA[dC]VK.D	Gly->Asp (G) [58.0055]	high
0.016	4.53E+04	R.TPEVTCVVVDVS[dW]EDPEVK.F	Ser->Asn (S) [27.0109]	high
0.024	6.70E+04	R.TPEVTCVVVDV[dW]EDPEVK.F	Ser->Asn (S) [27.0109]	low
0.172	6.35E+05	K.TTPPVLDSDGSFFL[Y]nK.L	Ser->Asn (S) [27.0109]	low
0.111	1.29E+05	K.DSTYSLSSTLTL[T]K.A	Ser->Asn (S) [27.0109]	high
0.179	2.00E+05	K.DSTYSLSSTLTL[T]K.A	Ser->Asn (S) [27.0109]	low
0.183	2.52E+05	K.HKVYACEVTHQ[dL]SSPVTK.S	Gly->Asp (G) [58.0055]	high
0.134	1.79E+05	K.HKVYACEVTHQ[dL]SSPVTK.S	Gly->Asp (G) [58.0055]	low

Table of SVA peptide spectrum matches (PSMs) for two LC-MS/MS analyses of the NIST reference mAb lot 31fb showing their relative abundance (XIC ratio %) relative to the wildtype. MS res. refers to high or low resolution for the MS2 spectra, all in CID mode. There are additional data fields not shown.

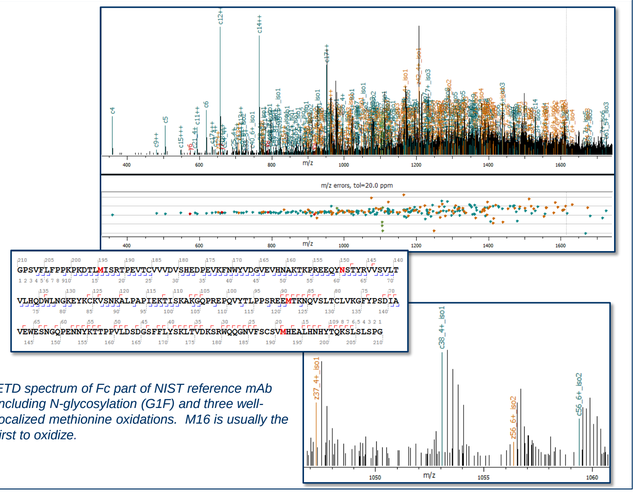
Glycosylation and Top-down analysis



This ETD spectrum shows a glycopeptide with G2F. ETD spectra show primarily peptide fragmentation, along with a few special peaks such as acetyl loss from HexNAc. HCD and CID also can produce successful identifications.



Glycopeptides with the same base peptide tend to co-elute. (Sialic acid residues cause a small predictable shift.) This behavior can help provide additional identifications of low- abundance glycopeptides in simple samples (Goldberg et al, 2007).



ETD spectrum of Fc part of NIST reference mAb including N-glycosylation (G1F) and three well-localized methionine oxidations. M16 is usually the first to oxidize.

Byologic® Features

- Convenient identification of sequence variants, modifications, and degradants down to concentrations of 0.1% of the unmodified form
- Visual comparison of MS1 and MS2 spectra of variant and unmodified forms
- Label-free quantification of variant relative to the unmodified sequence
- Compare multiple samples and multiple digestive enzymes
- Rapid rejection of false-positive identifications
- Exports figures and tables in variety of formats for reports and filings

Conclusions

To fully capitalize on the rapid advances in analytical instruments such as mass spectrometers, biopharmaceutical scientists require bioinformatics tools to process and manage the large influx of data. Byologic software brings together multiple sets of data to fully characterize therapeutic proteins or other proteins or protein complexes. Together with the Byonic and Byomap software by Protein Metrics, the software enables scientists to confidently characterize the product down to trace components identifying degradants, impurities, post-translational modifications, and sequence variants—including comparing across multiple samples. Byologic incorporates analysis techniques to transform into routine tasks these analyses that historically took expert analysts weeks.

An important aspect of the software is report generation, allowing analysts to rapidly produce reports and filing documents.

Not shown due to space limitations are Byomap results to address primary sequence analysis with peptide mapping requirements.

Acknowledgments

Protein Metrics gratefully acknowledges financial support from NIH grants GM100634 and GM103352. NIST Disclaimer: Certain commercial equipment, instruments, software or materials are identified in this presentation. Such identification does not imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the products identified are necessarily the best available for the purpose.

References

M. Bern, Y.J. Kil, C. Becker, "Byonic: advanced peptide and protein identification software." Curr. Protoc. Bioinformatics. 2012 Dec; Chapter 13:Unit13.20. PMID: 23255153 PMCID: PMC3545648

D. Goldberg, M. Bern, et al., "Automated N-Glycopeptide Identification Using a Combination of Single- and Tandem-MS." J. Proteome Res. 2007, 6, 3995-4005.